Revista Signos Fónicos
Vol 1 N° 1 enero/ junio 2015

REFLECTION PAPER

Mas Bayes, menos Chi-cuadrada
More Bayes, Less Chi-squared

Carlos Alfonso Mantilla Duarte

RESUMEN

Introducción: Desde que Neyman y Pearson desarrollaron la propuesta de Fisher para valorar la evidencia, χ^2 se consolidó como la prueba óptima para la toma de decisiones. En las últimas décadas, este enfoque frecuentista ha entrado en crisis permitiendo al enfoque bayesiano perfilarse como la mejor alternativa para esta toma de decisiones. **Métodos:** Estudio cualitativo de carácter explicativo. Mediante revisión documental se expone el funcionamiento del Enfoque Neyman-Pearson (prueba χ^2) y el Enfoque Bayesiano *(FB)*. Se presentaron dos casos donde se conocía la respuesta para analizar el comportamiento de las pruebas frente a la evidencia. **Resultados:** En el Caso 1, los resultados de la prueba χ^2 contradijeron la evidencia mientras que *FB* permitió seleccionar adecuadamente la respuesta. En el Caso 2, mediante χ^2 no fue posible encontrar solución para el Contraste de Hipótesis, el Enfoque Bayesiano, por el contrario, abordó de manera natural el problema. **Análisis:** El problema de χ^2 consiste en la imposición de un valor α para H_0 sin considerar el interés real del investigador. *FB* utiliza la información a priori que el usuario posee sobre el problema y las expectativas que el mismo tiene sobre la decisión que va a tomar. **Conclusiones:** Es preferible el uso de *FB* pues siempre permite encontrar la respuesta mientras χ^2 no funciona adecuadamente en todos los casos. **Palabras claves:** Chi-Cuadrada, Factor de Bayes, Razón de Verosimilitud.

ABSTRACT

Introduction: Since Neyman and Pearson developed Fisher's Works assesing evidence's strength, χ^2 it has been consolidated as the optimal test for decision making. In the last decades, this approach is in crisis allowing the Bayesian approach emerge as the best alternative for decision making. **Methods:** Qualitative study of explanatory nature. Through literature review explore the Neyman-Pearson approach χ^2 and the Bayesian approach *(FB)*. Two cases where the answer was known are used to analyze the behavior of the tests with the evidence presented. **Results:** In Case 1, χ^2 contradicted the evidence while *FB* admited the right answer. In Case 2, χ^2 can't found the anwer for Hypothesis Testing, Bayesian Approach, however, solved the problem naturally. **Analysis:** The χ^2 problem is the imposition of α value in H_0 excluding the real interest of the investigator. *FB* uses the prior information of user about the problem and the expectations that he has on the decision going to take. **Conclusions:** Is prefer to use *FB* because always permits to find the answer, while, χ^2 does not operate properly in all cases. **Keywords:** Bayes Factor, Chi-Square, Likelihood Ratio.

INTRODUCCIÓN

Desde que Neyman y Pearson desarrollaron la propuesta de Fisher sobre el método para valorar la fuerza de la evidencia, se ha definido a la prueba Chi-Cuadrada como aquella que ofrece resultados óptimos en las denominadas pruebas de significancia. Sin embargo, al tratarse siempre de una regla de decisión, la respuesta no debiera condicionarse únicamente a valores prefijados de sus probabilidades como se hace con el valor de α que se define por el tipo de error estadístico que se espera cometer. En la teoría de las decisiones existen alternativas más completas que aprovechan las creencias a priori y las combina con las expectativas a posteriori sobre la toma de una decisión, la cual se conoce como enfoque bayesiano y que se desarrolla a partir de la regla de Bayes sobre probabilidad condicional. Este documento expone, brevemente, los dos enfoques con el fin de resaltar las ventajas que tiene el hacer uso del Factor de Bayes en lugar de la prueba Chi-Cuadrada. Partiendo de un planteamiento teórico de ambos enfoques (enfoque bayesiano y enfoque de Neyman-Pearson) se pasa a la discusión sobre los resultados que ofrece cada una de las pruebas para, finalmente, concluir sobre cuál de ellas se adapta mejor a las condiciones de modelamiento de la realidad.

MÉTODOS

Por costumbre, los diferentes estudios científicos han incluido en sus análisis de datos una serie de pruebas que responden al condicionamiento del comportamiento mismo de los datos. En principio se supone que los datos han sido generados a través de un mecanismo aleatorio conocido que puede ser representado mediante un miembro F_θ de una familia paramétrica de distribuciones [(1)]. Bajo el supuesto paramétrico se diseñan las pruebas que buscan encontrar la hipótesis que mejor se ajusta a los datos. Para tales fines existen dos enfoques el Enfoque Bayesiano (Razón de Verosimilitud o Factor de Bayes) y enfoque de Neyman-Pearson (Chi-Cuadrada). La revisión documental sobre el planteamiento matemático de los enfoques permite entender cómo funciona cada uno de ellos.

Enfoque Neyman-Pearson.
Aunque el aporte de Ronald Fisher y el que diez años más tarde realizaron Jerzy Neyman y Egon Pearson eran en buena medida conceptualmente contradictorios entre sí; en la actualidad lo que se aplica es una especie de híbrido que toma del primero el valor p y del segundo la noción de que la tarea consiste en valorar si procede o no rechazar una hipótesis [(2)].
El planteamiento de la prueba Chi-Cuadrada se puede expresar como en la fórmula [1] [(3)]:

$$\chi^2{}_0 = \sum_i \frac{(F_i - p_i N)^2}{p_i N} \qquad y \qquad \chi^2{}_1 = \sum_i \frac{(F_i - q_i N)^2}{q_i N} \qquad [\ 1\ ]$$

Donde N representa el número de observaciones. Fi la frecuencia de ocurrencia de i, p la probabilidad de i y q su complemento. Se escoge, entonces, la Hipótesis con el menor χ^2, los subíndices representan las respectivas hipótesis o modelos.

Teorema de Bayes.
En general, una expresión extendida de la regla de Bayes puede ser expresada de la forma representada en la fórmula 2 [(4)]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} \qquad [\ 2\ ]$$

Donde $\bar{A}$ representa el complemento de A. Como se observa, la regla de Bayes produce probabilidades inversas. Además, incluye el condicionamiento de B cuando A ocurre $P(B|A)$. $P(A|B)$ representa la distribución *a posteriori* de A dado B y $P(A)$ la probabilidad a priori de A. Éste es el punto de partida para el planteamiento de la prueba basada en la regla de Bayes y que se denomina Factor de Bayes el cuál se expresa como:

$$\frac{P(x|H_0)}{P(x|H_1)} = FB \qquad [3]$$

Que se resuelve mediante una integral y el resultado obtenido indica si es el numerador quien supera al denominador o viceversa. Así, mediante el Factor de Bayes se compara la probabilidad de ambas Hipótesis y se indica cuál de ellas aventaja a la otra. Todo software estadístico, emplea una medida derivada del Factor de Bayes conocida como Razón de Verosimilitud que implementa el logaritmo natural con el fin de simplificar el cálculo del área bajo la curva de probabilidad en cada observación. Así se tiene:

$$ln\left(\frac{P(D|p)}{P(D|q)}\right) = \ln(LR) \qquad [4]$$

Con el enfoque Neyman-Pearson surge un problema para el investigador que no se presenta con el enfoque Bayesiano: debido al teorema del límite central la decisión se hace cada vez más difícil pues a medida que N aumenta las diferencias entre los modelos son casi imperceptibles. Además, la respuesta de Chi-Cuadrada depende, *"catastróficamente"*, de cómo hayan sido pareados los datos [3]. Bayes, por su parte, muestra la relación de verosimilitud existente entre las hipótesis o los modelos y señala aquella que aventaje a la otra.

RESULTADOS

Caso 1: Partida de dados [3]

Un ejemplo clásico es la determinación del ganador de una partida de dados con base en la frecuencia observada de las partidas y los resultados de una partida en particular:
Supóngase que las hipótesis son las siguientes:

$$p = \frac{1}{13*14}(1,1,1,1,1,1,1,1,1,1,1,1,1,13,13,13,13,13,13,13,13,13,13,13,13,13)$$

$$q = \frac{1}{13*14}(6,6,6,6,6,6,6,6,6,6,6,6,6,8,8,8,8,8,8,8,8,8,8,8,8,8)$$

En términos representación clásica de las hipótesis se tiene: $H_0{:}P=1$ $\qquad$ $H_1{:}P \neq 1$

Supóngase, ahora, que la frecuencia observada es:

$$F = (1,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1)$$

Por simple observación, los datos favorecen a p luego es de esperarse que los test arrojen los mismos resultados. ¿Pero, qué muestran las pruebas?

$$\chi_P^2 = \sum_i \frac{(F_i - p_iN)^2}{p_iN} = 12 \qquad \chi_Q^2 = \sum_i \frac{(F_i - q_iN)^2}{q_iN} = 9{,}29 \qquad ln\frac{P(D|p)}{P(D|q)} = 4{,}5$$

La respuesta obtenida mediante Chi-Cuadrada contradice lo esperado pues sugiere que el mejor modelo es Q. Está claro que el mejor modelo debiera ser P y no Q, tal como lo indican la intuición y la observación. De acuerdo al teorema de Bayes, los datos están en favor del modelo P en un 99%. Este es un ejemplo de comportamiento típico bajo el principio de incertidumbre.

Caso 2: Comparación de Distribuciones [5].
El siguiente ejemplo ilustra el comportamiento de los enfoques bajo condiciones atípicas:
Supóngase que se tiene un conjunto de observaciones i.i.d. de la forma $X_1, X_2, \ldots, X_n$ cuya distribución de los datos f es o una Normal o una Cauchy, entonces, se require contrastar las siguientes hipótesis:

$$H_0: f \sim N(\mu, \sigma^2) \qquad\qquad H_1: f \sim C(\mu, \sigma^2)$$

Esto, por sí mismo, es difícil de abordar desde una perspectiva clásica ya que no existe un estadístico de prueba natural ni una hipótesis nula convencional, incluso, fijando algunos parámetros como $\pi(\theta_0) = \frac{1}{\sigma^2}, \mu = 0, \sigma^2 = 1.$ Es decir, forzando los datos a una Distribución Normal Estándar.

Por su parte el enfoque Bayesiano aborda el problema de una manera más natural:

$$FB_{N-C} = \frac{\int C(\mu, \sigma^2)\pi_C(\theta)d\theta}{\int N(\mu, \sigma^2)\pi_N(\theta)d\theta}$$

Ya que la función de densidad de las distribuciones no depende de la probabilidad a priori no informativa el primer término de la integral es una constante por lo que el valor del Factor de Bayes es:

$$FB_{N-C} = \frac{C(\mu, \sigma^2)\int \pi_C(\theta)d\theta}{N(\mu, \sigma^2)\int \pi_c(\theta)d\theta}$$

El cuál dependerá única y exclusivamente de los valores de los parámetros de localización y escala del conjunto de datos. Si se supone $\pi(\theta_N) = \pi(\theta_C)$ y se adiciona la condición $\mu=0$ y $\sigma^2=1$, $FB_{N\text{-}C}=0{,}797884561$. Lo que sugiere que los datos corresponden a una Distribución Normal, como se planteó para buscar la solución desde el enfoque Neyman-Pearson.

Si se supone que $\pi(\theta_N) = \frac{1}{\sigma^2} FB = 2 \times 0{,}797884561$. Entonces, los datos corresponden a una Distribución Cauchy. Nótese que sin importar cuál sea el orden que se emplee para calcular el Factor de Bayes la evidencia apuntará a la Hipótesis que tenga la ventaja. Si se hubieran invertido los modelos el resultado sería: $FB_{C\text{-}N} = 1{,}253314137$. Fijando de nuevo los valores de los parámetros de localización y escala como se hizo anteriormente.

ANÁLISIS

¿Chi-Cuadrada o Bayes?

En ambos enfoques se parte del contraste entre la Hipótesis Nula (H_0) y la Hipótesis Alternativa (H_1). Como se sabe, se plantea la posibilidad de cometer dos tipos de errores: Error Tipo I (Falso Rechazo) y Error Tipo II (Falso No-Rechazo).

En el enfoque de Neyman-Pearson (Chi-Cuadrada) se escoge que la probabilidad del Error Tipo I sea α y no se da guía racional de cómo escogerle, siendo dictadas por la convención que $\alpha=0{,}1$; $0{,}05$ ó $0{,}01$ [4]; en contraste, el enfoque Bayesiano no fuerza la escogencia de α sino que define, para ello, una combinación de creencias a priori sobre las hipótesis y las pérdidas que encierra una decisión [4]. Clásicamente, el valor crítico de los estadísticos de prueba se interpretan como el valor mínimo de α para el cual se habría rechazado H_0. Es decir, se debe interpretar como la probabilidad que H_0 sea correcta contrariando, incluso, los mismos principios frecuentistas sobre *Hipótesis Correctas*.

Está claro que los dos enfoques no llevan a la misma respuesta. Se puede inferir un poco sobre el Factor de Bayes a fin de crear un criterio de comparación con Chi-Cuadrada. Así, mientras que chi-cuadrada muestra una dependencia cuadrática con F_i , (véase [1]), Bayes muestra una relación lineal con la misma variable (3):

$$ln\frac{P(x|H_0)}{P(x|H_1)} = \sum_i F_i ln\frac{p_i}{q_i} \qquad [5]$$

La pregunta que surge en este punto es: ¿cuál enfoque lleva a la respuesta correcta?

El problema del enfoque de Neyman-Pearson radica en que el investigador no busca valorar la probabilidad que tienen los datos bajo el supuesto de H_0 cierta sino lo contrario: la probabilidad de que H_0 sea válida partiendo del supuesto que se observaron los datos que dieron pie a las diferencias existentes entre la observación y su esperanza.

Como se observa en [1], la función obliga a escoger entre las dos hipótesis aquella con el valor más pequeño de χ^2, lo que lleva a excluir la otra hipótesis de manera inmediata. Bajo este enfoque sólo es posible decidir con base en la probabilidad de la hipótesis escogida previa imposición del valor de α.

En [3] se observa un fenómeno más interesante: además de permitir explicar la probabilidad de la Hipótesis Nula se puede explicar la probabilidad de la Hipótesis Alternativa. Este enfoque, además de proveer una interpretación más adecuada del valor de α, requiere que el usuario suministre parte del insumo: ¿cuáles son las pérdidas de una decisión incorrecta y qué se conoce a priori sobre H_0? Esta condición sugiere que Chi-Cuadrada puede llevar a conclusiones erróneas en situaciones muy comunes como ocurre con el Caso 1.

Ahora bien, ¿Qué ocurre con el Caso 2, es decir, donde las hipótesis que se contrastan no son equivalentes? ¿Cuál es el planteamiento desde el enfoque Neyman-Pearson? ¿Cómo se aborda el problema según Bayes?

Para el caso 2, ambos enfoques requieren información inicial para determinar los valores de p y de q para lo cual se empleó la probabilidad no informativa $\pi(\mu, \sigma^2) = \frac{1}{\sigma^2}$ (6). Esto, permitió intentar darle solución mediante χ^2, así, se fija $p = \frac{1}{\sigma^2}$ y $q = 1 - p$. Ya que el modelo Q no es complementario para P este planteamiento es erróneo por lo que el enfoque Neyman-Pearson empieza a tener dificultades para contrastar las hipótesis. Supóngase, entonces, que se permite fijar los parámetros de localización y escala de la siguiente manera: $\mu=0$ y $\sigma^2=1$. Como se dijo, forzando los datos a una Distribución Normal Estándar. Sin embargo, χ^2 sigue sin poder dar solución. Bayes, en cambio, resuelve el problema sin mucho esfuerzo y, cuando se añaden las condiciones seleccionadas para χ^2, se obtiene una solución numérica fácil de interpretar debido a que el Factor de Bayes es una *razón de verosimilitud.*

Ahora bien, ¿y si el problema es más complejo? Supóngase, ahora, que se tiene una situación en la cual es necesario decidir por una de tres situaciones: intervenir inmediatamente, esperar un plazo e intervenir, no intervenir. En términos de test de significancia se tendría:

H_0:*Intervenir Inmediatamente* H_1:*Esperar e Intervenir* H_2:*No Intervenir*

Cómo abordar esta situación desde el Enfoque Neyman-Person y cómo desde el Enfoque Bayesiano. Esta situación corresponde a reglas de decisión. De nuevo, la prueba χ^2 aporta poco a la discusión ya que no permite contrastar simultáneamente más de dos hipótesis, además de no poder realizar la comparación de las hipótesis de manera natural. El Factor de Bayes aunque con un algoritmo más complejo que el expuesto en este documento puede ayudar a tomar una decisión. Claro, dependerá el resultado de la información *a priori* que posea el usuario y de las expectativas que el mismo tenga sobre el problema. Pero se llega a una conclusión.

CONCLUSIONES

• En la práctica, es casi seguro que Chi-Cuadrada funciona bien. Sin embargo, en casos atípicos la prueba lleva a conclusiones equivocadas. Incluso, al momento de moldear la realidad simple, Chi-Cuadrada puede llevar a respuestas erróneas. El Factor de Bayes ofrece mejores alternativas al momento de decidir sobre las hipótesis pues mediante esta prueba se recopila información que el usuario tiene sobre los datos y las combina con las expectativas pues mezcla tanto la probabilidad de la hipótesis de prueba como la probabilidad de su hipótesis complementaria (en términos estadísticos: la probabilidad a posteriori se define por la probabilidad a priori). En otras palabras: χ^2 casi siempre funciona, FB siempre funciona. ¿Por qué emplear una prueba que casi siempre funciona si existe una que siempre funciona?

• En la vida real, las decisiones se basan en la información previa y su combinación con las expectativas. Chi-Cuadrada deja de lado las expectativas condicionando la respuesta a la imposición de un valor fijo de α. El enfoque bayesiano no fuerza la escogencia de α sino que define una combinación de creencias a priori sobre las expectativas que hay respecto a la toma de una decisión. Además, el Factor de Bayes ofrece ventajas sobre el enfoque de Neyman-Pearson ya que no supone que los datos obedecen a una distribución fija sino deja que la naturaleza del fenómeno sea quien determine su comportamiento.

• Muchas situaciones reales son difíciles de modelar mediante el Enfoque Neyman-Pearson mientras que el Enfoque Bayesiano las aborda naturalmente. Más aún cuando se trata de alternativas que no son complementarias. El modelamiento bayesiano es aplicable a todas las ciencias, principalmente a las ciencias de la salud y a las ciencias económicas ya que las decisiones en estos campos dependen objetivamente del conocimiento previo que se posee y de las expectativas que hay sobre la selección de las alternativas. La incertidumbre sobre las decisiones siempre ha existido. Desde el s. XVIII existe un modelamiento matemático que simplifica la decisión que se abandonó desde los años 30's del s. XX para dar paso a un modelo que ignora el comportamiento impredecible de la naturaleza aportando más incertidumbre a los problemas.

REFERENCIAS

1. Zamar RH. Estimación Robusta. Estadística Española. 1994; 36(137): p. 327 - 387.

2. Servizo Galego de Saúde. Análisis Bayesiano. [Online].; 2006 [cited 2015 Enero. Available from:

http://www.sergas.es/.
3. MacKay DJC. Bayes or Chi-squared? Or does it not matter? [Online].; 2005 [cited 2015 Enero. Available from: http://www.inference.eng.cam.ac.uk/is/.

4. Pericchi Guerra LR. Análisis de Decisión, Inferencia y Predicción Estadística Bayesiana Caracas: Centro de Estadística y Software Matemático (CESMa) y Dpto. de Cómputo Científico y Estadística Universidad Simón Bolívar; 1999.

5. Pericchi Guerra LR. Bayes Factors: The Need to Replace P-Values [Memorias X Coloquio de Estadística]. Medellín; 2014.

6. Berger JO, Pericchi LR. The Intrinsic Bayes Factor for Model Selection and Prediction. Journal of the American Statistical Association. 1996 Marzo; 91(433): p. 109-122.